

ABSTRACT**METHOD AND SYSTEM FOR THE AUTOMATIC SEGMENTATION OF AN AUDIO
STREAM INTO SEMANTIC OR SYNTACTIC UNITS**

A digitized speech signal (600) is input to an F0 (fundamental frequency) processor that computes (610) a continuous F0 data from the speech signal. By the criterion voicing state transition (voiced/unvoiced transitions) the speech signal is presegmented (620) into segments. For each segment (630) it is evaluated (640) whether F0 is defined or not defined i.e. whether F0 is ON or OFF. In case of F0 = OFF a candidate segment boundary is assumed as described above and, starting from that boundary, prosodic features are computed (650). The feature values are input into a classification tree and each candidate segment is classified thereby revealing, as a result, the existence or non-existence of a semantic or syntactic speech unit.